

(43) Date of A Publication 20.06.2001

(21) Application No 9929520.6

(22) Date of Filing 15.12.1999

(71) Applicant(s)

Graeme John Proudler  
5 Touchstone Avenue, Stoke Gifford, BRISTOL,  
BS34 8XQ, United Kingdom

(72) Inventor(s)

Graeme John Proudler

(74) Agent and/or Address for Service

Kenneth Stanley Targett  
48 Meadowsweet Way, HORTON HEATH, Hampshire,  
SO50 7PD, United Kingdom

(51) INT CL<sup>7</sup>

G10K 11/175

(52) UK CL (Edition S )

H4R RPX

(56) Documents Cited

US 5749324 A

US 4464119 A

US 4202293 A

US 3823691 A

(58) Field of Search

UK CL (Edition R ) A1M MCA , A5K K6 , G5G G7 , H4J  
JGA , H4R RPX RSN RSRE RSX  
INT CL<sup>7</sup> A01K 15/02 , A61F 5/58 , G10K 11/16 11/175  
11/178 , H04R 3/00 3/02 27/00  
online: EPODOC,WPI,JAPIO

(54) Abstract Title

Audio processing, e.g. for discouraging vocalisation or the production of complex sounds

(57) Various audio processing methods and apparatus are described for discouraging vocalisation or the production of complex sounds. In one method, a signal is created from undesirable incident ambient audio (106) and is processed (at 103), possibly under the influence of controls, and converted to output audio which is broadcast (from 105) so as to mix with the undesirable incident ambient audio. The processing for at least the majority of control settings causes oscillatory ambient audio in common ambient environments. In another method, a signal is created which may be generated from detected ambient audio (210), or may be a predetermined or pseudo-random signal (203, 204). Ambient audio is used to selectively enable output of audio (from 209) produced from that created signal, and the output audio is broadcast so as to mix with ambient audio. The output audio (from 209) is broadcast in timed bursts (e.g. to interrupt an aggressive speaker). Stable positive feedback may be promoted.

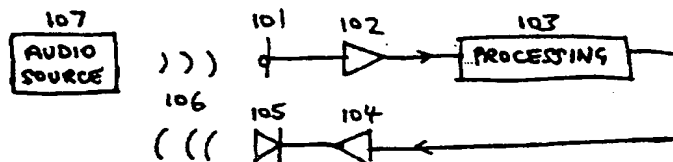


Figure 1

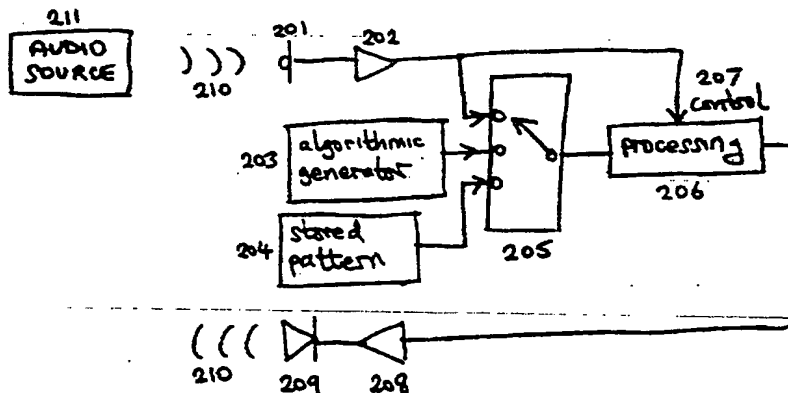


Figure 2

1/3

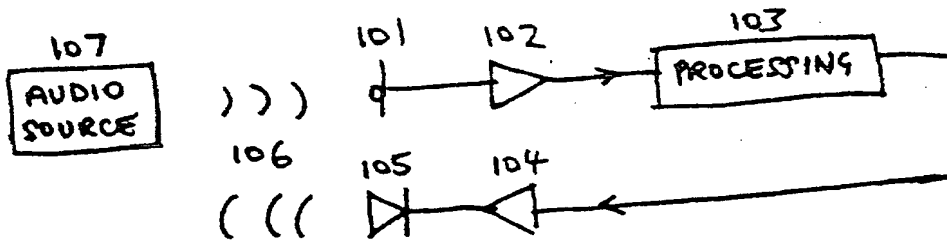


Figure 1

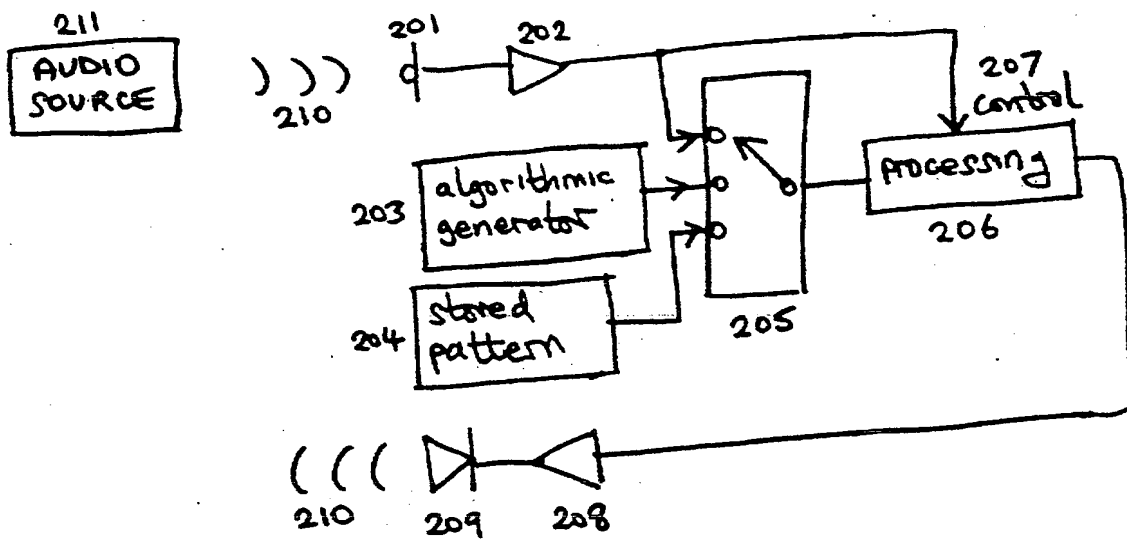


Figure 2

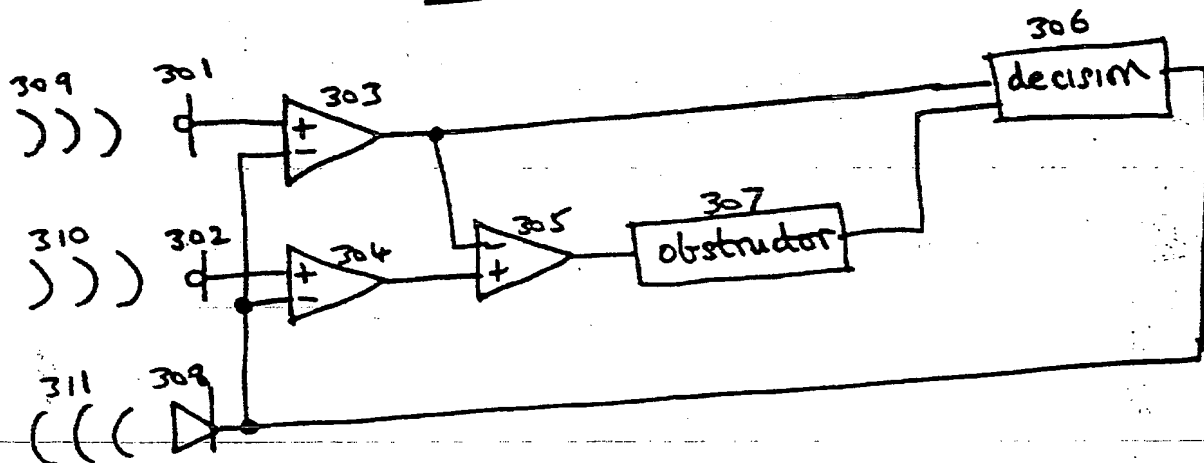


Figure 3

Best Available Copy

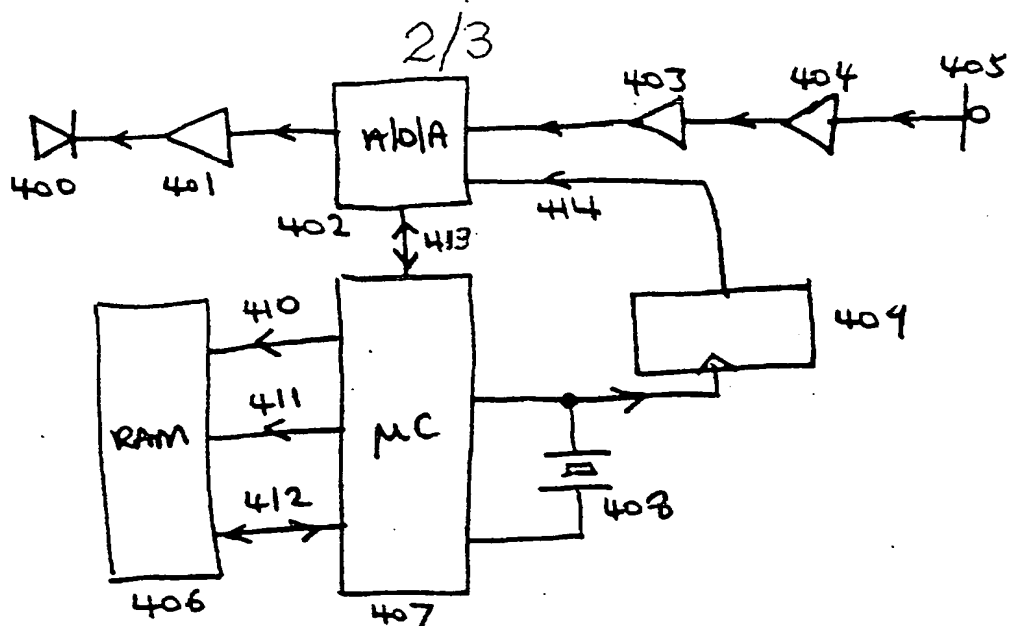


Figure 4

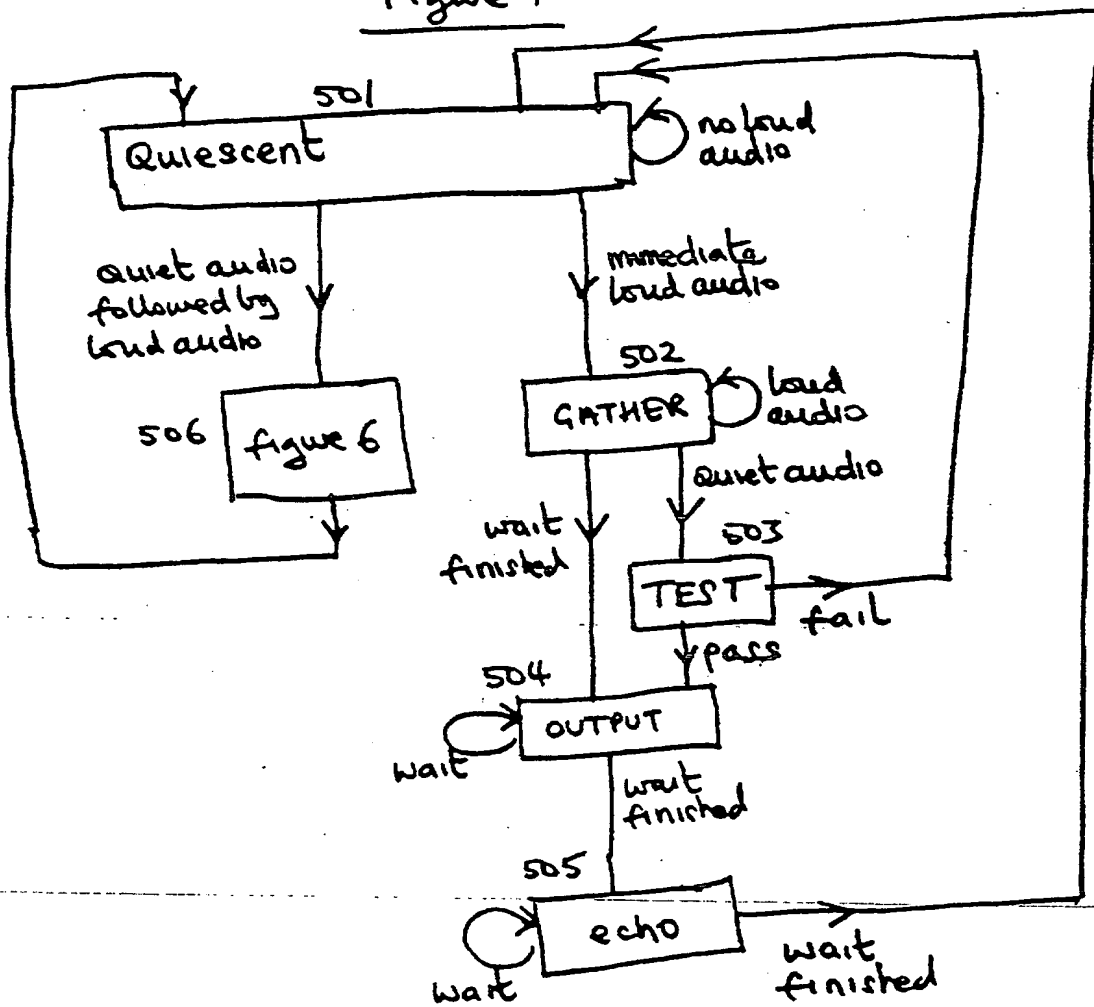


Figure 5

3/3

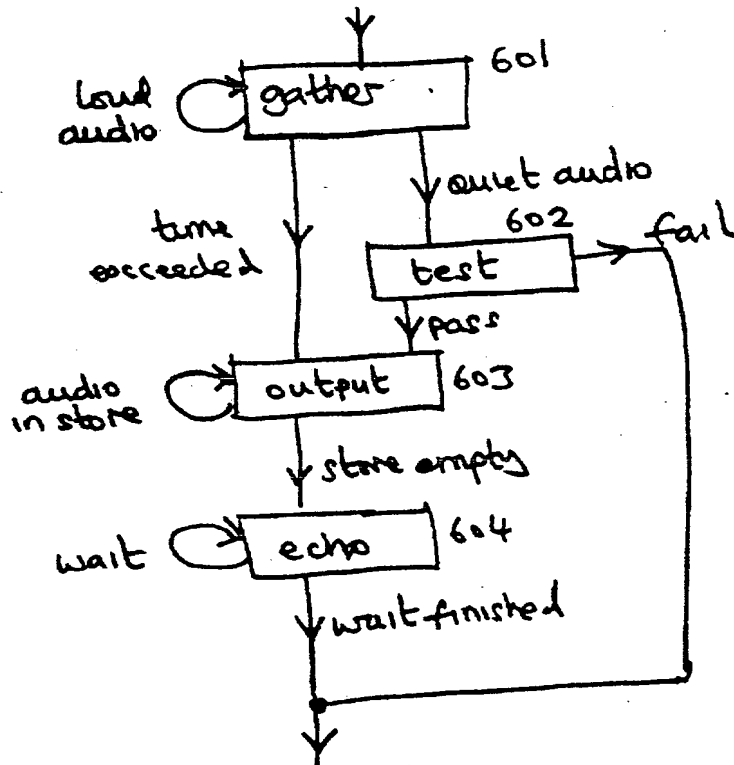


Figure 6

TITLE

Audio Processing, e.g. for Discouraging Vocalisation or the Production of Complex Sounds

DESCRIPTION

This invention relates to audio processing methods and apparatus, particularly (but not exclusively) for use in discouraging vocalisation or the production of complex sounds.

In this invention, the term 'vocalisation' includes not only speech but also other sounds or noises uttered by both human beings and also animals, and the term 'complex sounds' includes other sounds and noises such as music whether generated live or being a replay of a recording. The term 'ambient audio' implies an ensemble of sounds from a larger volume compared to that of 'localised audio', which implies far fewer sounds (perhaps just one specific sound) whose source is in the immediate region of a sensor. Ambient audio is not necessarily produced for the express purpose of detection by an audio sensor, while localised audio is often produced just for that purpose. Detection of ambient audio generally requires much greater amplifier sensitivity than detection of localised audio.

Often vocalisation or other complex sounds are unwelcome. Situations may occur, for example, during the course of employment involving contact with members of the public, or control of unruly individuals. An employee, for example at a social security office or a football ground turnstile or a railway station, may feel threatened by vocalisation, or be required to regain control of a situation but be unable or unwilling to apply direct force. Such threatening situations reduce the effectiveness of the employee and can cause job-related stress. It is therefore desirable to provide support for employees in such situations. There are presently few if any methods of providing such assistance: the employee must wait out the situation or try to verbally interrupt the unwanted vocalisation.

The present invention is concerned with discouraging such vocalisation and/or production of other ambient audio. Some methods described herein may be said to 'interfere' with undesirable spoken words, since they produce output ambient audio at the same time as the undesirable spoken words. Other methods may be said to 'interrupt' a speaker, since they reflect spoken words back to the speaker just after the end of an undesirable spoken word, in the same way that a person would normally interrupt another person.

A first aspect of the present invention provides for the creation of a signal from undesirable incident ambient audio, processing of that signal possibly under the influence of

controls, conversion of that processed signal to output audio, and broadcasting of that output audio so as to mix with the undesirable incident ambient audio, where the processing for at least the majority of control settings causes oscillatory ambient audio in common ambient environments.

- 5           The processing may cause continuous oscillation, or the oscillation may be repeatedly switched on and off in bursts.

A second aspect of the present invention provides for the creation of a signal, using ambient audio to selectively enable output audio produced from that signal, and broadcasting that output audio so as to mix with ambient audio.

- 10           The signal may be created from undesirable incident ambient audio, and/or from a source independent of incident audio, such as a white noise generator, a coloured-noise generator, or an oscillatory-signal generator, or combinations thereof. When the signal is created from incident undesired ambient audio, that incident undesired ambient audio may be used almost immediately or may be noticeably delayed.

- 15           The production of output audio may be dependent upon some or all of the following methods and events: inspecting desirable ambient audio to determine the characteristics of desirable audio that distinguish loud and quiet desirable audio; determining the presence and/or absence of loud desirable ambient audio; the presence of quiet desirable ambient audio; the presence of loud desirable ambient audio, inspecting undesirable ambient audio to determine the  
20           characteristics of undesirable audio that distinguish loud and quiet undesirable audio; determining the presence and/or absence of loud undesirable ambient audio; the presence of quiet undesirable ambient audio; the presence of loud undesirable ambient audio.

- An intentional delay may be provided between the detection of loud ambient audio and production of output audio. In one mode the output audio is produced before the end of loud  
25           ambient audio. In another mode the output audio is produced just after the end of loud ambient audio.

- Determining the presence and/or absence of loud ambient audio may involve some or all of the following: ignoring incident ambient audio while broadcasting output audio, ignoring  
incident ambient audio for a first time after broadcasting output audio; conditionally ignoring  
30           incident audio for a second time after broadcasting output audio, where the second time is longer than the first time.

Once broadcasting has started, it may continue for a time independent of ambient audio or may continue for a time dependent on the detection of quiet audio.

The method of the first and/or second aspect of the invention may be combined with a further method, such that desired audio may be broadcast instead of output audio produced according to the first or second methods. This has the effect of providing a conventional  
5 loudhailer when desired audio is detected.

Specific embodiments of the present invention will now be described, purely by way of example, with reference to the accompanying drawings, in which:

- Figure 1        schematically illustrates a first example of a method according to the present  
10                invention;
- Figure 2        schematically illustrates a second example of a method according to the present  
                 invention;
- Figure 3        schematically illustrates a third example of a method according to the present  
                 invention;
- 15    Figure 4        is a block diagram of an apparatus for performing the first and second  
                 examples; and

Figure 5 and 6 are a state diagram to illustrate the operation of the apparatus of Figure 4.

Figure 1 illustrates the first method. Ambient audio 106 is converted by microphone  
101 into a signal that is amplified to usable levels by preamplifier 102. The signal is processed  
20 by a processing block 103, amplified by a power amplifier 104, and broadcast by loudspeaker  
105. The broadcast audio mixes with audio from an undesirable audio source 107 to form the  
ambient audio 106.

The processing block 103 may or may not have external controls. It is capable of  
creating positive feedback between the microphone 101 and the loudspeaker 105 in essentially  
25 all ambient conditions. The actual nature of the ambient audio 106 will depend on the acoustic  
environment and the audio produced by the audio source 107. The processing block 103 may  
operate to produce continuous positive feedback, or the positive feedback may be repeatedly  
switched on and off. A non-interfering signal such as silence is produced when positive  
feedback is switched off. A typical burst duration would be 200ms.

The processing block 103 may be implemented in many ways that will be apparent, in the light of this specification, to those skilled in the art of signal processing.

5 A simple example of suitable processing in processing block 103 to produce interfering audio is automatic-gain-control without an activation threshold. The method is to inspect samples of incident ambient audio over recent time (perhaps a few tens of milliseconds) in order to determine the peak amplitudes of incident ambient audio. Even if the ambient audio environment is initially quiet, noise inherent in all circuitry will usually provide an irreducible level of background signal. An amplification factor is then calculated, such that those samples with peak amplitudes are amplified to the maximum desirable amplitude. This amplification factor is then applied to all samples before they are converted to audio. If such amplification would cause a new sample to have an amplitude greater than the maximum desirable value, the amplification factor is reduced so the new sample is amplified to the maximum desirable amplitude. The effect of this processing is ambient audio of an oscillatory nature, provided that the implementation has sufficient loop gain to compensate for the loss between the output transducer 105 and the input transducer 101. If positive feedback is to be repeatedly switched on and off, the processing block 103 outputs a signal generated *via* feedback when switched on, and a signal that represents silence when switched off.

Figure 2 is a general illustration of the main elements of the second method. Ambient audio 210 is converted by the microphone 201 into a signal that is amplified to a usable level by preamplifier 202. The signal is connected to a combiner/switch 205 and to the control input 207 of a processing block 206. Also connected to the switch 205 is an algorithmic generator 203 that produces a signal according to an algorithm. Also connected to the switch is a pattern generator 204 that produces a signal according to a stored pattern, which may be an artificially created pattern or a recording of a real audio signal. The switch connects some combination of the output of the preamplifier 202, the algorithmic generator 203, and the pattern generator 204 to the signal input of the processing block 206. The output of the preamplifier 202 controls the processing block 206 *via* its control input 207 to produce an output signal that is amplified by power amplifier 208 and broadcast by loudspeaker 209. The broadcast audio mixes with audio from an undesirable audio source 211 to form the ambient audio 210.

30 The output of combiner/switch 205 could therefore include a component from a pseudo-random source (such as that produced by algorithmic generator 203) or from a stored repetitive waveform source (such as that produced by the pattern generator 204). Such sources are well known *per se*.



If the output of combiner/switch 205 includes incident ambient audio derived from microphone 201, the processing block 206 may act to encourage ambient audio oscillation or may act to prevent ambient audio oscillation. Oscillation will occur if processing block 206 introduces sufficient loop gain. Oscillation may be prevented if processing block 206 ignores input audio while output audio is being broadcast. Then the apparatus may be said to operate in a 'record-or-replay' mode, since the processing block 206 gathers incident audio or outputs audio, but never does both simultaneously. Oscillation may also be prevented if processing block 206 uses 'echo cancellation' techniques to remove broadcast output audio from an input signal that includes both new incident audio and broadcast output audio. Then the apparatus may be said to operate in a 'record-while-replay' mode, since output audio can be broadcast while new incident audio is being gathered. Such 'echo cancellation' techniques are well known *per se* to one skilled in the art, and will not be mentioned further here except to note that such techniques require 'training' to learn the characteristics of the path between the output and input of the processing block 206. Such training necessarily requires the production of output audio in the absence of significant new incident audio. Sometimes this is done by deliberately producing a specific training signal. Training may be done while processing block 206 executes a 'record-or-replay' method. (This training method assumes that output audio is loud enough to dominate new ambient incident audio.)

The combiner/switch 205 and processing block 206 operate to produce a signal which represents output audio that discourages the production of ambient audio. In tests, interfering with spoken words by broadcasting a shrieking, shrill, oscillatory sound was found to be very assertive and effective, while interrupting speech by reflecting a spoken word (at the end of that spoken word) was more polite but less effective. Generally, output audio could be noise, or an alarm sound, a shrieking sound, or a delayed version of undesirable ambient audio, or any other audio that is found to be effective for the desired purpose.

The processing block 206 examines the signal presented at control input 207 so that loud ambient audio and quiet ambient audio may be differentiated and detected. This may be done in many ways, which will be apparent, in the light of this specification, to those skilled in the art of audio processing.

The type of output produced by processing block 206 depends on the presence of loud ambient audio, detected *via* the signal at control input 207. If loud audio has been detected, the processing block outputs a signal that represents the audio that will obstruct production of ambient audio. Otherwise, the processing block 206 outputs a signal that represents silence or some other audio that will not obstruct production of ambient audio. An output signal is

produced by processing block 206 from its input signal after the detection of loud ambient audio via control input 207.

5 A first mode of operation of the arrangement shown in Figure 2 'interferes' with spoken words, preferably before they have finished. In such a first mode, processing block 206 outputs interfering audio before the end of that loud ambient audio, in order to interfere with the loud ambient audio. A delay between detection of loud ambient audio and output of an interfering signal is provided to enable determination of the characteristics of the control signal that indicate loudness and quietness. This delay also enables detection of loud ambient audio via control input 207 (depending on the method used). The delay is also used to determine the recent peak amplitudes of the input signal to the processing block 206, which may be temporarily stored for future use in automatic-gain-control. The delay also enables the processing block 206 to reject signals at control input 207 that arise from bursts of ambient noise and unwanted echoes of previous output audio, as will be explained later.

15 The action of the processing block 206 when producing an interfering output signal is to amplify its input signal into an output signal that produces output audio with consistently loud mean output amplitude. If the output signal of the combiner/switch 205 is independent of the preamplifier 202, the output signal from the processing block 206 is simply amplified. If the output signal of the combiner/switch 205 is dependent on the preamplifier 202, the output signal from the processing block 206 is adjusted according to stored peak amplitudes of the signal input and new peak amplitudes of the signal input. Methods of applying automatic-gain-control will be apparent, in the light of this specification, to those skilled in the art of audio processing. Preferably the interfering output signal is controlled so that it does not overdrive the power amplifier 208 or the loudspeaker 209. Once production of an output signal from processing block 206 has started, it continues for a preset time. While the processing block is producing an interfering output signal, the processing block 206 assumes that it cannot differentiate between signals at its control input 207 that were caused by original ambient audio and those that were caused by output audio. So the processing block 206 freezes detailed interpretation of its control input 207. The processing block 206 also freezes detailed interpretation of its signal input, except as previously noted when preamplifier 202 contributes to the signal source.

30 A second mode of operation of the arrangement shown in Figure 2 'interrupts' speech during gaps in that speech. In such a second mode, processing block 206 starts the output of interrupting audio just after a break in the incident undesired audio. This mode reflects essentially whole spoken words back to a speaker, either almost immediately after that word was finished, or a short time later. The combiner/switch 205 is operated to produce its output

from the preamplifier 201. Processing block 206 acts to prevent oscillation and false triggering by applying either the 'record-or-replay' or 'record-while-replay' methods described above to both its signal input and its control input 207, to isolate genuinely new ambient audio.

5 If the processing block 206 is executing the 'record-or-replay' method, such isolation is achieved simply by the act of ignoring input signals while producing interrupting output audio. The overall effect is that processing block 206 detects new loud ambient audio, stores that audio until it becomes quiet, replays that stored audio and simultaneously ignores ambient audio, and then returns to searching for new loud ambient audio.

10 If the processing block 206 is executing the 'record-while-replay' method, such isolation is achieved by subtracting a delayed version of output audio from input signals. The overall effect is that every piece of new loud ambient audio is delayed until it becomes quiet and is then replayed. The processing block 206 isolates new ambient audio from its input signal and stores it in temporary memory. The processing block 206 isolates new ambient audio at its control input 207 and detects the start of new loud ambient audio. When new isolated quiet  
15 ambient audio is detected *via* control input 207 after new isolated loud ambient audio, the processing block 206 outputs the stored input signal from temporary memory, from the start of the new isolated loud ambient audio to the start of the new isolated quiet ambient audio.

In both 'record-or-replay' and 'record-while-replay' modes:

1. Automatic-gain-control is applied to maintain a uniformly high mean level of output  
20 audio.
2. A minimum amount of audio is stored before processing block 206 produces output audio. Otherwise, the stored audio is discarded. This is to eliminate activation of the second mode by spurious bursts of noise.
3. Processing block 206 automatically starts replay of stored audio when a preset  
25 maximum amount of audio has been stored. This is to eliminate lockup of the second mode in the presence of continuously loud new ambient audio.

During the first mode and second mode executing 'record-or-replay' (but not 'record-while-replay'), when the production of output signal ceases, the processing block 206 rejects both its input and signals at control input 207 for a short time to allow the amplitude of ambient  
30 echoes of output audio to decay below the level that will be interpreted as loud audio. The processing block 206 then conditionally accepts larger signals at control input 207 as being caused by new original loud ambient audio provided that the signal is large for longer than a certain time. Obviously this time must be shorter than the delay between the detection of loud

ambient audio *via* control input 207 and the decision to create an output signal. The result of such unconditional and conditional rejection is that production of new obstructing audio caused by old output audio is much reduced, if not eliminated. The louder (earlier) loud echoes of output audio are simply ignored. The quieter (later) echoes are rejected if they are not masked  
5 by new loud ambient incident audio.

There are many variations on the methods described above. They could be used on their own, or could be combined with other known or obvious methods. For example, if audio is quiet for a period and then loud audio is detected, the method of delaying whole spoken words could be used. Otherwise, if there appears to be a substantial amount of loud audio, an  
10 interfering method could be used. This has the overall effect of interrupting a speaker if there is a small amount of loud ambient audio present, and interfering with speech if there is a large amount of loud ambient audio present. This combined method is useful because interrupting is a modest form of assertion and sufficient to dissuade some but not all individuals from speaking, while interfering is a more robust form of assertion, and dissuades more individuals from  
15 speaking. Using this algorithm, the method will continue interrupting if interruption is effective. Otherwise, it will use interference.

Another variation is to activate the method depending of the time of day, the relative occurrence of loud ambient audio, and so on.

Another variation is to add at least one sensor that detects desirable audio. The  
20 detection of loud audio at that sensor takes precedence over the detection of undesired audio and causes desired audio to be broadcast from a, or the, loudspeaker instead of obstructing audio. Normally such desirable audio will be localised audio, such as words spoken directly into a microphone, instead of ambient audio. This is because it must be possible to distinguish desired audio from undesired ambient audio. It is, however, possible for desired audio to  
25 originate at a distant source. Its general form is illustrated in Figure 3, where an audio sensor 301 produces an input signal from ambient desired audio 309 and another audio sensor 302 produces an input signal from ambient undesired audio 310. A loudspeaker 308 is driven by the output of decision circuit 306. Obstructor circuit 307 produces an obstructing signal using one of the methods previously described. The overall principal of the variation is that decision  
30 circuit 306 outputs a signal derived from audio sensor 301 when desired audio is active, and otherwise outputs an obstructing signal from obstructor circuit 307. The output signal is subtracted from the desired input and also from the undesired input using subtractors 303 and 304, such that any trace of the output signal is at an acceptably low level. It may also be

necessary to remove the clean desired signal from the clean undesired signal using a subtractor 305, such that any trace of the clean allowed signal is at an acceptably low level.

The general case may be simplified in several ways, including:

1. If the 'record-or-replay' method is in use, there is no need to subtract output audio from undesired input audio. This eliminates subtractor 304.
2. In the special case where the desired audio comes from a significantly different direction to undesired audio, the use of a directional microphone pointed towards the undesired audio will pick up the undesired audio but not the desired audio, thus eliminating the stage of removing the desired signal from the undesired signal. This eliminates subtractor 305.
3. In the special case where the output audio comes from a significantly different direction to the desired input or the undesired input, the use of directional microphones pointed away from the output audio will not pickup the output signal, thus eliminating the stage of removing the output signal from the desired signal and from the undesired signal. This eliminates subtractors 303 and 304.
4. In the special case where the desired signal is produced using a non-audio transducer, such as throat microphone, the desired signal will not include output signal, thus eliminating the stage of removing output audio from desired audio. This eliminates subtractor 303.
5. In the special case where desired audio is much louder than undesired audio, the amplitude of input audio from a single sensor can be compared to a threshold, and input audio processed as desired audio when above that threshold, or processed as undesired audio when below that threshold.

Figure 4 illustrates the preferred physical architecture. An electret microphone-insert 405 converts ambient audio into an electrical signal that is magnified by amplifiers 404 (such as the National Semiconductor LM358 set for a gain of 2) and 403 (such as the National Semiconductor LM386 bypassed for maximum gain). The output of amplifier 403 is the audio input to a codec 402 (such as the Texas Instruments TCM320AC36). The codec 402 is driven by control signals generated by the microcontroller 407 (such as a Microchip PIC16C64). The codec 402 converts the incident analogue audio to digital and compresses it to an 8 bit word (using  $\mu$ law coding in this example).

The microcontroller 407 controls the codec 402 via reset, data, clock and sync signals such that the codec sends the compressed data to the microcontroller 407, and performs manipulation of the data according to the program stored inside the microcontroller 407. The microcontroller 407 has insufficient internal temporary memory, and therefore uses the RAM 406 (8k  $\times$  8 industry standard type 6264) to store the compressed data samples. The

Best Available Copy

microcontroller 407 produces address signals 410 and control signals 411 to drive the RAM 406. The microcontroller 407 exchanges data with the RAM 406 via data signals 412. When the microcontroller has finished its processing, it sends a compressed digital version of the output audio to the codec 402 using signals 413. The codec converts the digital data to an analogue waveform that is amplified by the power amplifier 401 (such as Analogue Devices SSM2211), that drives the loudspeaker 400 (such as a 1.5W loudspeaker).

The microcontroller derives its timebase from the crystal 408 (preferably 20MHz). The crystal 408 also drives a counter 409 (such as the industry standard HC4024) that produces a reference clock 414 for the codec 402.

10 If the method involves the storage of ambient audio, the microcontroller 407 continually drives the RAM 406 so that compressed input data is continually written to the RAM. New data overwrites the oldest data when the RAM is full. The microcontroller is also continually inspecting input data to detect contiguous loud audio. There are many ways of determining when loud audio is present, all of which will be apparent, in the light of this specification, to one skilled in the art. In a prototype, time was divided into arbitrary contiguous intervals of 15 20ms or so, the peak value in each interval was noted, and the last nine peak values recorded in a FIFO. An upper threshold is set to half the median value in the peak FIFO. When the input amplitude exceeds the upper threshold, a 20ms or so retriggerable 'upper-monostable' is set. A lower threshold is set to an eighth of the median value in the peak FIFO. When the input 20 amplitude exceeds the lower threshold, a 20ms or so retriggerable 'lower-monostable' is set. If the prototype's state is 'audio absent', the state changes to 'audio present' when the 'upper-monostable' is active. If the prototype's state is 'audio present', the state remains as 'audio present' as long as the 'lower-monostable' is active. The actual start of contiguous audio is taken to be 20ms or so before the state changes to 'audio present'. The actual end of contiguous 25 audio is taken to be 20ms or so after the state changed to 'audio absent', when the state has been 'audio absent' for 80ms or so. It will be appreciated that this is just one method of determining the presence or absence of spoken words, that the values quoted here can be varied, and that there are other methods.

Figure 5 is an illustration of a state-machine that is implemented as a program in the 30 microcontroller in the preferred implementation. The program in the microcontroller examines the samples representing incident ambient audio.

In Figure 5, new loud incident ambient audio is examined for loud audio during the QUIESCENT state 501, and the characteristics of quiet audio are updated. When loud audio is detected, the state changes.

5 If the program spends more than a short time (120ms in the prototype) in the QUIESCENT state 501, the program executes the method 506, where entire spoken words are replayed as soon as they have finished. (This is illustrated in Figure 6.) Then the program returns to the QUIESCENT state 501.

10 If the program spends less than a short time (120ms in the prototype) in the QUIESCENT state 501, the state changes to GATHER state 502. In GATHER state 502, the amplitude of detected audio is examined so as to temporarily record the peak levels of the audio, and the characteristics of loud audio are updated. If audio becomes quiet, the state changes from the GATHER state 502 to TEST state 503.

15 In TEST state 503, the time since the broadcast of output audio is measured, and the duration of the loud audio is examined. If the time since broadcast of output audio is too short (the prototype used a duration of 140ms), or the duration of the loud audio is too short (the prototype used a duration of 180ms), the audio is rejected and the state returns to QUIESCENT state 501. Otherwise, the state changes to OUTPUT state 504.

In GATHER state 502, if the time spent reaches a limit (the prototype used a duration of 180ms), the state changes to OUTPUT state 504.

20 In the OUTPUT state 504, audio is generated from a signal, and is broadcast. When the time spent in OUTPUT state 504 reaches a limit (the prototype used a duration of 180ms), the state changes to ECHO state 505.

25 In the ECHO state 505, all ambient audio is ignored. When the time spent in ECHO state 505 reaches a limit (the prototype used a duration of 20ms), the state returns to QUIESCENT state 501.

The preferred implementation uses incident audio as the signal that is converted to audio and broadcast. The audio sample that has just been gathered is amplified by an automatic gain control to produce a consistently loud mean output amplitude without clipping. The microcontroller does this by noting the maximum sample amplitude during the GATHER state 502, and amplifying all samples by the same amount so that the maximum sample amplitude during replay is the peak desired value. If feedback causes larger input samples that would be clipped by this process, the amount of amplification is reduced so as to avoid clipping.

An alternative implementation could use a signal derived from an algorithmic generator. One example is the use of a pseudo-random generator to produce apparently random noise. (A description of pseudo-random generators is in 'Pseudo Random Sequences and Arrays' - MacWilliams and Sloane, proc. IEEE vol. 64 #12, December 1976.) A suitable  
5 polynomial is  $[x^{15} + x + 1]$ , since it has few taps but has a cycle length of a few seconds when incremented once per sample period. The contents of the generator could be repeatedly exclusive-ORed with audio samples during the start of the GATHER state 502 to provide a variable start position when the time comes to provide output audio, provided that steps are taken to detect the all-zero lockup state and exit it. An audio sample could be produced from  
10 the generator by incrementing it every sample period. The six least significant bits of the generator are used to produce a varying audio output. Four bits are used as the amplitude part of a  $\mu$ law sample, another bit as the least significant bit of the segment value of that sample, and another bit as the sign bit. The two most significant bits in the segment value should be set to 1, to ensure a large amplitude output. This produces 'white' noise audio, which may be  
15 acceptable for interrupting certain speakers.

Another alternative implementation could use a signal derived from a primitive pattern stored in non-volatile memory. At each sample period, a successive value of the pattern is converted to audio. When the end of the pattern is reached, the method cycles back to using the start of the pattern, and the process repeats. Such patterns (such as sine wave, or more complex  
20 cyclic signals) may be generated by algorithms, while others (such as a stored version of actual positive audio feedback) may be stored versions of actual audio signals .

Method 506 (where entire spoken words are replayed as soon as they have finished) is illustrated in Figure 6. In GATHER state 601, the amplitude of detected audio is examined so as to temporarily record the peak levels of the audio, the characteristics of loud audio are  
25 updated, and detected audio is temporarily stored. If audio becomes quiet, the state changes from GATHER state 601 to TEST state 602.

In TEST state 602, the time since the broadcast of output audio is measured, and the duration of the loud audio is examined. If the time since broadcast of output audio is too short (the prototype used a duration of 140ms), or the duration of the loud audio is too short (the  
30 prototype used a duration of 180ms), the audio is rejected and the state returns to QUIESCENT state 501 shown in figure 5. Otherwise, the state changes to OUTPUT state 603.

In GATHER state 601, if the time spent reaches a limit (the prototype used a duration of 400ms), the state changes to OUTPUT state 603.

Best Available Copy



In the OUTPUT state 603, audio is replayed from the store, automatic-gain-control is applied and audio is broadcast. When the store is empty, the state changes to ECHO state 604.

In the ECHO state 604, all ambient audio is ignored. When the time spent in ECHO state 604 reaches a limit (the prototype used a duration of 20ms), the state returns to  
5 QUIESCENT state 501 shown in Figure 5.

It should be noted that the embodiments of the invention have been described above purely by way of example and that many modifications and developments may be made thereto within the scope of the present invention.

CLAIMS

1. An audio processing method, for example for discouraging vocalisation or the production of complex sounds, the method comprising the steps of:  
  
detecting incident ambient audio to produce a detected signal; and  
  
producing output audio from an output signal and broadcasting the output audio so as to  
5 mix with the incident ambient audio, the output audio being broadcast in bursts timed in  
dependence upon the detected signal.
2. A method as claimed in claim 1, wherein the presence of incident audio is ignored for a predetermined time after each such burst of output audio.
3. A method as claimed in claim 1 or 2, further comprising the steps, in the case of a  
10 burst of the ambient audio, of:  
  
determining whether the duration of the burst of ambient audio is less than a  
predetermined time; and  
  
if so, disabling the broadcasting of such a burst of output audio in response to that burst  
of ambient audio.
- 15 4. A method as claimed in any preceding claim, wherein the content of the output signal is  
produced at least in part from the content of the detected signal.
5. A method as claimed in claim 4, wherein the content of the output signal is produced at  
least in part from the substantially current content of the detected signal.
6. A method as claimed in claim 4, wherein the content of the output signal is produced at  
20 least in part from delayed content of the detected signal.

Best Available Copy

7. A method as claimed in claim 6, further comprising the steps, in the case of a burst of the incident ambient audio, of:

detecting the start of the burst of the incident ambient audio; and

commencing such a burst of the output audio a predetermined time after the detected start of the incident burst.

5

8. A method as claimed in claim 6, further comprising the steps, in the case of a burst of the incident ambient audio, of:

detecting the end of the burst of the incident ambient audio; and

commencing such a burst of the output audio a predetermined time after the detected end of the incident burst.

10

9. A method as claimed in claim 6, further comprising the steps, in the case of a burst of the incident ambient audio, of:

detecting the start of the burst of the incident ambient audio;

determining whether or not the detected start is more than a first predetermined time after the end of the previous burst of output audio; and,

15

if so:

detecting the end of the burst of the incident ambient audio; and

commencing such a burst of the output audio a second predetermined time after the detected end of the incident burst; but,

20 if not:

commencing such a burst of the output audio a third predetermined time after the detected start of the incident burst.

10. A method as claimed in any of claims 4 to 9, further comprising the step of processing the detected signal to produce the output signal so as promote stable positive feedback.

25 11. A method as claimed in claim 10, wherein the processing step includes the steps of:

Best Available Copy

determining the peak value of the detected signal in a recent period; and  
amplifying the detected signal with a gain generally inversely proportional to the that  
peak value to produce the output signal.

12. A method as claimed in any of claims 4 to 9, further comprising the step of processing  
5 the detected signal to produce the output signal so as prevent positive feedback.

13. A method as claimed in any preceding claim, wherein the content of the output signal is  
produced at least in part from a source independent of the incident ambient audio.

14. An audio processing method, for example for discouraging vocalisation or the  
production of complex sounds, the method comprising the steps of:

10 detecting incident ambient audio to produce a detected signal;  
processing the detected signal to produce a processed signal; and  
producing output audio from the processed signal and broadcasting the output audio so  
as to mix with the incident ambient audio to form a feedback loop;  
wherein the processing step is controlled so as promote stable positive feedback in the  
15 feedback loop.

15. A method as claimed in claim 14, wherein the processing step includes the steps of:  
determining the peak value of the detected signal in a recent period; and  
amplifying the detected signal with a gain generally inversely proportional to that peak  
value to produce the processed signal.

20 16. A method as claimed in claim 14 or 15, further comprising the step of intermittently  
disabling the feedback loop so as to produce bursts of positive feedback.

Best Available Copy

17. A method as claimed in claim 16, wherein the period of each burst is less than two seconds, more preferably less than one second, more preferably less than 500 ms, and more preferably about 200 ms.
18. A method as claimed in claim 16 or 17, wherein no audio is broadcast between  
5 successive bursts of positive feedback.
19. A method as claimed in any preceding claim, further comprising the steps of:  
detecting further audio to produce a further detected signal; and  
modifying the output audio when the existence of such further audio is detected.
20. A method as claimed in claim 19, wherein, when the existence of such further audio is  
10 detected, the output audio is produced from the further detected signal.
21. An audio processing method, for example for discouraging vocalisation or the production of complex sounds, substantially as described with reference to the drawings.
22. An audio processing apparatus adapted to perform the method of any preceding claim.



INVESTOR IN PEOPLE

Application No: GB 9929520.6  
Claims searched: 1-13 at least

Examiner: Martyn Dixon  
Date of search: 13 June 2000

## Patents Act 1977 Search Report under Section 17

### Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

UK Cl (Ed.R): H4R (RSRE,RSN,RSX,RPX); H4J (JGA); A5K (K6); G5G (G7); A1M (MCA)

Int Cl (Ed.7): H04R (3/00,3/02,27/00); G10K (11/16,11/175,11/178); A01K (15/02); A61F (5/58)

Other: Online: EPODOC, WPI, JAPIO

### Documents considered to be relevant:

| Category | Identity of document and relevant passage                         | Relevant to claims |
|----------|---|--------------------|
| X        | US 5749324 A (Moore) see col 6, lines 48-64 and col 8, lines 7-26 | 1,3,13,22          |
| A        | US 4464119 A (Vilgrube <i>et al</i> )                             |                    |
| X        | US 4202293 A (Tri-Tronics) see e.g. col 4, lines 47-58            | 1-3,13,22          |
| X        | US 3823691 A (Morgan) see the whole document                      | 1,4-6,13,22        |

## Best Available Copy

|   |   |   |  |
|---|---|---|--|
| X | Document indicating lack of novelty or inventive step   | A | Document indicating technological background and/or state of the art   |
| Y | Document indicating lack of inventive step if combined with one or more other documents of same category. | P | Document published on or after the declared priority date but before the filing date of this invention.          |
| & | Member of the same patent family  | E | Patent document published on or after, but with priority date earlier than, the filing date of this application. |